Abstract and Keywords

Historically it has been easier to focus on measuring and describing differences between groups of people rather than try to describe the dynamic ways that individuals change. Dynamical systems are mathematical models that aim to describe how constructs change over time. Frequently these models are continuous time models; models that try to capture the function that underlies a set of observations. This chapter introduces the concept of a dynamical system and of continuous time models. Two methods are introduced for the fitting of continuous time models to observed data: one using the approximate discrete model for a first-order autoregressive model and the second using a method of estimating latent derivatives for a second-order autoregressive model.

Keywords: Dynamical System(s), Dynamic System(s), Continuous Time, Discrete Time, Differential Equation Model(s)(ing), Derivative Estimation, Approximate Discrete Model

Introduction

Short time-scale, intraindividual variability is often hard to model in the social sciences. It is of little surprise that researchers chose to begin by modeling more macroscopic, interindividually features first (e.g., mean differences among groups). Increasingly, perhaps because of an interest in causal relationships, the questions being asked by some researchers have moved from questions about group differences to questions of change. In answering how constructs change, emphasis has been placed on more macroscopic features first (e.g., long-term linear or quadratic change). But these long-term changes are not the essence of living—they are the product of countless minutes, hours, and days
of incremental changes. The analysis of individuals on shorter time-scales is still frequently ignored because the data often appear so complex that they could be confused with random variation.

But to understand the essence of how individuals grow and adapt in response to their environment, the analysis of very short time-scales is necessary. Sitting at a wedding, one might ask oneself: “Will this couple be happy 10 years from now?” One could look for macroscopic predictors of success—perhaps similarities in political disposition—but this would only begin to unravel the picture. John Gottman’s “5 to 1 ratio” of positive to negative interactions, however, suggests that the key to positive marriages may be built in daily events and moment-to-moment interactions. Many useful predictors that don’t change over the course of many days have been identified. The usefulness of these predictors, however, may be due to their ability to reflect what is occurring at much shorter time scales. For example, people of similar political dispositions would seem likely to have a few less topics on which they could have negative interactions, which in turn might lead couples with similar dispositions to be closer on average to the 5 to 1 ratio than couples with differing political dispositions. If long-term outcomes are the product of many smaller events and decisions, studies that examine long-term development may average over rich, informative variability, just as averaging over a group can average over the rich differences between individuals. But even if the analysis of individuals at short time-scales is essential to understanding how people develop, how could one go about conceptualizing and modeling such change?

This chapter introduces tools for the conceptualizing and modeling of nonlinear change through the concepts of dynamical systems and continuous time models. There are many excellent introductions to dynamical systems (e.g., Smith & Thelen, 2003; Thelen & Smith, 2006); the present chapter blends some introductory concepts with two tools that have been developed for empirical research on relatively short time series (< 100 observations). This chapter begins by discussing concepts related to dynamical systems and continuous time modeling. Two methods for analysis of time series are then presented: the approximate discrete model (Bergstrom, 1988; Oud, 2007) and a method that estimates latent derivatives (Boker & Nesselroade, 2002). Sample code for each of the methods is provided in the appendices and on the website of the author.
Dynamical Systems: The Concept

Many statistics are based on the analysis of the consistent parts of a system: the steady-states of constructs, trait-like constructs, constructs that have high test-retest reliability. To speak of dynamical systems, on the other hand, is to convey an interest in that which changes, and that which is not constant, that which is inherently unstable. The first concept to address is that of a system. A system is all of the interrelated elements in the domain being studied. Systems could consist of one person or several people, a single construct or several constructs. A dynamic system is a system where the elements change over time. This differs from a dynamical system, which is a mathematical model of a dynamic system. The translation of dynamic systems—systems that change—into the language of mathematics is the goal of dynamical systems.

Dynamical systems are broadly categorized into one of two classes: linear dynamical systems and nonlinear dynamical systems. The distinguishing feature of these classes is how the predictors are combined. If the predictors in all equations are multiplied only by constants and added up, as is typically how predictors are entered into regression equations, then the dynamical system is linear. If the predictors are multiplied with each other, or there are terms such as the exponent of a predictor, then the dynamical system is nonlinear. The classification of dynamical systems is by the linearity or nonlinearity of the equations—not of the resulting trajectory. Many seemingly complex, nonlinear trajectories can be described using linear equations—that is, linear dynamical systems.

In linear dynamical systems, changes in the predictors result in proportional changes in the dependent variables. Systems where there are non-proportional changes—for example, sudden transitions between states are indicative of a nonlinear dynamical system. One example is axon firing. The firing of an axon is often described as an all-or-none event, where a very small change in input could result in no change to the axon or alternatively could tip an axon into firing. Such threshold effects are one example of a feature that may convey the need for a nonlinear dynamical system. The present chapter focuses on introducing linear dynamical systems, but there are many resources for readers interested in nonlinear dynamical systems (e.g., Kaplan & Glass, 1995; Strogatz, 1994; Thompson & Stewart, 1986).
One remarkable thing about dynamical systems is that even linear systems can produce intricate change over time. The time series in Figure 19.1 are examples of the complexity that can be observed with a linear system; note that there is no error in these time series, they are the result of linear, regression-like equations. The possibility of describing change—particularly complex changes in observed variables—stirs the excitement of many fields, and so literature on dynamical systems can be found in a variety of natural and social sciences (e.g., physics, chemistry, biology, medicine, physiology, psychology, economics, etc.). The study of dynamical systems tends to leave many researchers with an indelible excitement because it presents fascinating and seemingly counter intuitive ideas, for example: the idea that extremely complex changes over time do not have to be the result of complex processes or models. In trying to explain the daily fluctuations of a complex system—perhaps the construct of stress—one could seek a model that tries to include predictors to explain every hill and valley that occurs—a complex model (many parameters) to explain seemingly complex process; no doubt the reader can think of a dozen factors or more that contribute to stress, all of which might need to be included to model stress. From a dynamical systems perspective, however, many of the observed changes may be attributable to the dynamics of stress itself—perhaps the self-regulation of stress or the simple interactions of stress with another part of the system (e.g., negative affect; Montpetit, Bergeman, Deboeck, Tiberio, & Boker, 2010). Even if stress and negative affect tended to fluctuate with a weekly schedule and there were absolutely no external influences (a simple model), one could see complex trajectories such as those in Figure 19.1 if they affect each other; complex trajectories can arise even without a complicated set of external influences. Fundamentally these mathematical models are intricately bound to a quest for simplicity and functionality that conveys an elegance that elicits the emotions typically reserved for things of overwhelming beauty. People got so lost in the beauty of dynamical systems that in the 1980s and 90s that much of dynamical systems theory was co-opted by *Chaos Theory*; Chaos Theory was introduced by many writers as a new way of thinking, the herald of a true scientific revolution. Chaos Theory and dynamical systems are, however, older than many writers convey and can be traced at least as far back as the work of mathematician Henri Poincaré (1854–1912).
The Language of Dynamical Systems

The essence of dynamical systems are models of systems that change. The mathematical models used therefore must be able to describe how a construct is changing with respect to time. In mathematics, describing the change in one variable with respect to another variable is usually accomplished using derivatives. Introduction to derivatives occurs early in many education systems when students first learn about the slope of the line—that is, “rise over run” or the change in $y$ divided by the change in $x$. Although the name may not have been used, the linear slope is also called the first derivative of $y$ with respect to $x$. Although the word derivative seems to be fear-provoking for many people, the first derivative sneaks into our daily lives in many ways. Innocuous letters go by like “mph” without announcing that miles per hours is just shorthand for the first derivative of position—that is, a change in position (measured in miles) with respect to a change in time (hours).

The second derivative builds on the first by expressing the change in the first derivative with respect to time. In our cars, the second derivative expresses the change in mph per unit time—that is, it expresses how quickly one is changing the speed (first derivative) of the car whether through acceleration or deceleration (e.g. braking). One other derivative that is commonly used is the zeroth derivative. This derivative expresses the position of the car at some time. The terms zeroth, first, and second refer to the order of the derivative. Derivatives are by no means limited to the second order (acceleration) and one can think about higher order derivatives—the third order, for example, would convey information about how quickly acceleration is changing with time.

In the social sciences there is frequent talk about the zeroth derivative—the level of a person’s construct at some time. Many constructs change with time, however, and we could then used two pieces of information to described a person—the level of a person’s construct at some time, and the speed at which the construct is changing (imagine the slope of a line). Many forms of change are not well described with lines. Rather than continuing up or down at some constant speed, frequently there will be one or more changes in speed, such that the line curves in some manner. This curve represents a change in speed with respect to time and indicates a non—zero second derivative. We could then imagine that at any given time, a person has a particular score on some construct, but that construct is changing with some speed (linear slope), and simultaneously the speed that the person is changing may be increasing or decreasing (accelerating or decelerating). Equivalently, we could describe the construct of a person at any given time and how that construct is changing using the zeroth, first, and second derivatives.

For people interested in change, the question then becomes: At any given time, what predicts whether a person’s scores are increasing or decreasing (positive or negative first derivative)? accelerating or decelerating (positive or negative second derivative)? One can imagine trying to find predictors of the first and/or second derivative(s) to address
what might relate to how a person’s scores change. Any model with a derivative, on either side of the equation, is called a differential equation model. Differential equation models can be used to express changes in a system of constructs; they can be used to describe the relationships between the current state of a construct and how that construct is changing, how a construct changes in response to the level of a another construct, or how a construct changes is response to changes in other constructs.
Attractors and Self-Regulation

The idea of relationships between the state of a construct and how it changes brings up one dynamical systems concept that may be particularly relevant to the social sciences; that concept is self-regulation. In the process of trying to describe all the small changes in a construct we could think about using a large number of predictors to try to explain every small change observed. Alternatively, we might find that we can model constructs in terms of a relationship between current states and change, such that people perturbed from their typical state might have a tendency to change so as to return to their typical state or equilibrium—they might self-regulate. The concept of self-regulation is a natural one for social scientists, as many constructs seem to exhibit homeostasis. In dynamical systems, one will often read of the concept of an attractor. An attractor is a state or set of states around which a dynamical system will fluctuate or converge. Many people will imagine the idea of a marble moving in a bowl, with the bottom of the bowl being an attractor; one can also imagine weaker and stronger attractors, much like a bowl with shallower or steeper sides. The idea of an attractor seems to map well onto the ideas of homeostasis, equilibrium, one’s “typical” state, or one’s trait.

Systems are not limited to one attractor but can have many attractors. Waddington’s epigenetic landscape (Fig. 19.2) is a famous visualization of a dynamical system with a changing number of attractors (Waddington, 1957). Waddington’s figure involves imagining a marble rolling toward the viewer as time progresses. Initially the ball can easily waver to the right and left in a single attractor. As the ball rolls forward, it may find itself in either one of two attractor states. The ball can still vary to the right and left, and with enough of a push could even surmount the hilly obstacle in the center and vary around the alternative attractor. As time goes on, the number of attractor states changes and the various attractors differ in their depth. Some attractors will consist of deep wells, allowing relatively little variation in the ball’s movements and few possibilities to change attractors; other attractors will be more shallow, allowing for more variation and more possibility of switching attractors.

Waddington’s interests were in the process of tissue differentiation. A chicken begins as a set of undifferentiated cells that over time becomes increasingly differentiated so as to produce organs, muscles, bones, and so forth. Early in the process, if one removes several of the undifferentiated cells, then the chicken will still develop normally—one does not expect that by removing a few undifferentiated cells that one can produce a boneless chicken. Early on the cells are gathered around a single, shallow attractor that allows for change to occur easily. Over time the attractor wells deepen, such that late in the process, cells have a much harder (if not close to impossible) time surmounting the attractor walls so to become another type of tissue; muscle and bone will not change to replace a missing kidney.
For psychological constructs, we might imagine a slightly different landscape. We might imagine that early in life, people are predisposed to certain traits but that those traits are not specifically defined. Through interactions with their environment, the landscape changes. Perhaps the valley of the attractor becomes deeper than people were initially predisposed, leading people to vary less in a construct and suggesting that there are periods that traits would be more precisely measured. However, it is well known that even mighty rivers, over time, can begin to carve a new path. Perhaps similarly, these constructs might be altered over time such that the position of the valley switches. Figure 19.3 is one way to imagine the effects of events “pulling” on the landscape (perspective from under the landscape), changing its shape. The conceptualization gets more complex, as one can imagine events (such as a traumatic life event) either changing the shape of the landscape or moving people from one attractor to another. Furthermore, people could differ in the depth of their attractor(s); with some people being easily perturbed within or between shallow valleys and some people varying only within very deep valleys.

*Figure 19.2* Waddington’s epigenetic landscape (1957). As the ball rolls down the valley (progression of time), small differences in initial conditions lead the ball to different paths (attractors). The application of the correct force, however, could push the ball from one attractor to another. One can also imagine the ball varying within an attractor basin.
The attractors described so far have primarily been point attractors or combinations of a few point attractors. There are several types of attractors, of which the point attractor is just one. For example, one could think about a point repeller. As its name implies, this is a state from which a system diverges rather than converges. It is a point of instability from which a system will depart, given a small amount of energy.

Rather than imagining a ball in a bowl (attractor) or a valley, imagine balancing a marble on top of a larger ball. With a bowl, the marble is attracted to the lowest part of a basin. On the other hand, if one balances a marble on top of another ball, then it is unlikely to remain there for very long because a small amount of energy will knock the marble from its current state. Like the valley example, repellers can have very steep or very shallow walls—they can differ in their strength. Imagine balancing a marble on something with very steep sides (perhaps the head of a pin) versus something with less steep sides (perhaps a basketball). There are also many different shapes of attractors; frequently they are divided into categories that define the attractors as either a point attractor, a periodic attractor, or a chaotic attractor. Attractors become increasingly complex beyond a point attractor, but all are based on the fundamental idea of the state or states to which a system converges (or for repellers—diverges).

### Discrete and Continuous Time

This chapter began by introducing the ideas of systems of dynamic variables, models of change, derivatives, and differential equation models to introduce the mathematical tools and language of dynamical systems. The ideas of attractors and self-regulation gave a sampling of the concepts and metaphors often used by people working with dynamical systems. This section begins to move the reader toward the application of a dynamical system to data. When considering a changing construct, it seems almost natural to consider change with respect to time—that is, time is often selected as the baseline variable with respect to which other variables change. Although time is often not the variable of central focus, differing modeling techniques treat time very differently. One
The primary way that models can differ is the treatment of time as either discrete or continuous.

The expression of changes in variables using derivatives and differential equation models was coming into its full swing by the time Newton and Leibniz were telling the world about calculus, but it was not until almost 100 to 150 years later (late 1700s to mid-1800s) that differential equation models started to be approximated using difference equations (Lakshmikantham & Trigiante, 2002). Difference equations are expressions of the relationships between consecutively observed values—for example, that the observation at some time is equal to the previous observation times a constant plus some error ($x_t = Ax_{t-1} + \epsilon e$) with no attempt to model the system at times between $t - 1$ and $t$. These difference equations, which treat time as if it is discrete, are much easier to work with than differential equations, which treat time as if it is continuous.

For the physical sciences, where measurements could be rapidly and frequently made (relative to the social sciences) with relatively little error and with little change occurring between subsequent measurements, difference equations produced very reasonable approximations. This was particularly convenient, as it would not be until the early and mid-twentieth century before tools were becoming available for the fitting of stochastic differential equation models (essentially differential equation models with random errors). In the mean time, however, the use of difference equations spread to many other fields, including the social sciences. By the mid—twentieth century there was trouble on the horizon. As Bergstrom described the history of econometric analysis: “At the time when Bartlett’s paper [1946] was published, econometricians were becoming increasingly aware of the problems created by interaction between variables within the unit observation period.” (Bergstrom, 1988)

What econometricians like Bergstrom realized was described by Bartlett (1946)

> “The discrete time nature of our observations in many economic and other time series does not reflect any lack of continuity in the underlying series. Thus theoretically it should often prove more fundamental to eliminate this imposed artificiality. An unemployment index does not cease to exist between readings, nor does Yule’s pendulum cease to swing.”

Barlett’s concern was the way that discrete time models treated the data. Difference equations treat subsequent observations as if they were caused by previous observations, ignoring the interaction of constructs between observations, as in Figure 19.4a. If we imagine this figure to represent two psychological variables (perhaps the monthly interactions between mother and child), then discrete time approaches can yield models that are strange from a theoretical perspective. For example, Figure 19.4a suggests that a child’s current score is related to his/her score a month prior and the mother’s score a month prior, but that the mother and child have not continued to affect each other daily over the course of the month. Discrete time models treat the mother-child interaction as
if it ceases to exist between measurements, that they interact at a few, discrete moments rather than continuously. But mothers and children do “not cease to exist between readings....”

The primary consequence of the seemingly innocent choice to use a discrete time model is that all of the results depend on the specific rate data were sampled; this issue has been addressed in the psychological literature in the context of longitudinal mediation (Gollob & Reichardt, 1987, 1991). There are several problems that follow. Three have the potential to lead to serious conflicts in the literature. First, the selection of the best model will depend on the observation interval. For example, if one collects monthly data conforming to a process that depends on the previous two observations ($x_t = A_1x_{t-1} + A_2x_{t-2} + \epsilon$), then by collecting quarterly data, one will find that these data will satisfy a completely different model (a moving average model; Bergstrom, 1988). Second, the relationships between variables will change depending on the frequency of observation. It has been shown that the effect of one variable on another can change from being positive, to non existent, to negative depending on the frequency that one makes observations (Oud, 2007). Finally, the magnitude of effects can also vary with sampling rate. Asking whether construct A has more of an effect on construct B, or vice versa, is a question whose answer will depend on the sampling rate—for some sample rates, A may have more effect on B (than vice versa), and for other sampling rates, B may have more of an effect on A (Oud, 2010).

These consequences have the potential to create important conflicts in the literature, merely due to researchers using different sampling rates and the use of discrete time methods (e.g., cross-lagged panel models). One way to think about these conflicts is to think about discrete time methods as a microscope focused on only a single sampling interval; analyzing monthly data with a discrete time model, we only get answers regarding a monthly interval but we can’t be sure that results extrapolate to other sampling intervals (e.g., bimonthly measurements or biweekly measurements). Aside from these conflicts, discrete time models often can have problems handling unequally spaced observations or missing data; in these cases, one must either start estimating differing parameters for the two differing time intervals (two different parameters in Fig. 19.5b) or find ways to create equal intervals such as the use of latent variables (e.g., Fig. 19.5c)
The mismatch with theory, the overwhelming dependance of results on the sampling rate, the limited interpretation of results, and the problems with unequal intervals (or missing data) would seem to be serious deterrents to the use of discrete time methods. Yet, despite the work of many economists, much of economic research is still based on discrete time methods, as is the case in much of the other social sciences. No doubt, in part, the reason is that it is easy to specify discrete time models—models where one observation is regressed on the previous observation are abundant in psychology. It may also be that researchers are not widely familiar with the weaknesses of discrete time models. Although more challenging to implement, continuous time models can be used with exactly the same data as discrete time models and surmount the weaknesses discussed.

One can think of continuous time models as trying to understand the functions that underlie the data, like the lines shown in Figure 19.6. Rather than saying one observation was caused by another, each observation is just a sample of many possible observations from an ongoing process. As studies sample individual subjects from some larger population, for a continuous time model the observations are a sample of individual observations (dark circles) from some larger population of possible observations (gray circles). With two constructs, there are two such functions. When one starts to think about how these functions affect each other, one can imagine changes in one construct at any moment being related to small, ongoing changes in the other construct—that is, changes in one construct continuously affect the other, like the many arrows in Fig. 19.4b, rather than just at specific discrete times. In addition, thinking about the observations as samples from a larger whole means missing observations and unequally spaced intervals are frequently less of an issue for these methods, much like how we are not usually concerned about people that aren’t sampled, unless there is something systematic about the people that weren’t sampled.
The following sections move from theory to application by demonstrating two methods for applying two different continuous time models. The first model is a first-order differential equation model, for which an autoregressive process is a solution in discrete time; an example of a discrete time version of such a model is shown in Figure 19.5a. This continuous time model will be fit using the Approximate Discrete Model (Bergstrom, 1988; Oud, 2007, 2010). The second continuous time

---

**Figure 19.5** Structural equation models of (a) a discrete time first-order autoregressive model, (b) the same model with a missing observation, (c) the same model with a way to estimate the autoregressive parameter, (d) a discrete time second-order autoregressive model, and (e) a way to think about the approximate discrete model, including instantaneous recursive paths (constraints not represented).

**Figure 19.6** Continuous time models aim to understand the continuous function underlying a set of observations (continuous line). As one is trying to understand the underlying function, these models are applicable with both equally (a) and unequally (b) spaced observations (black circles). The gray circles represent a larger population of possible observations.
model will be a second-order differential equation model—that of a damped linear
oscillator or pendulum. This model can be fit using a second-order autoregressive model
in discrete time, such as in Figure 19.5d. This continuous time model will be fit using
Latent Differential Equations (Boker, Neale, & Rausch, 2004). It should be noted that
either method can be used to fit either model, with differing advantages and
disadvantages; the two methods are demonstrated with differing models so as to expose
the reader to more options both for models and for methods. Moderate familiarity with
structural equation modeling (SEM) is assumed for the sections that follow; readers
unfamiliar with SEM might consider first reading Chapter 15 (this volume) or perusing an
introductory book (e.g., Kline, 2004).

First-Order Differential Equation Model

Recall that discrete time models were used as an alternative to continuous time models as
a method to simplify estimation. One very common discrete time model is the first-order
autoregressive model,

\[ x_t = A \Delta x_{t-\Delta} + w_t \]  

(1)

where \( x_t \) is the observed value at some time \( t \), \( x_{t-\Delta} \) is the observed value at some prior
time \( t - \Delta \), \( \Delta \) is the time between subsequent observations, and \( w \) is an error (sometimes
called innovation) that is often assumed to be normally distributed with mean of zero and
with values that are independent. The relationship between \( x_t \) and \( x_{t-\Delta} \) is captured in the
parameter \( A_\Delta \); the A subscript is used here to indicate that this value of \( A \) will depend on
the sampling rate—that is, the time between subsequent observations. This equation can
be shown to be a solution for the continuous time equation

\[ \frac{dx_t}{dt} = Ax_t + G \frac{dw_t}{dt}, \]  

(2)

which states that the first derivative of a variable at some time \( \frac{dx_t}{dt} \) is equal to the zeroth
derivative at some time \( x_t \) multiplied by a constant \( A \), plus error \( G \frac{dw_t}{dt} \). The second part
of the error term, \( \frac{dw_t}{dt} \), is a continuous time equivalent of \( w_t \) called the Wiener process.
Like \( w_t \), when integrated over some period of time this process produces independent,
normally distributed values, with a mean of zero. The first part of the error term, \( G \), is
included because the Wiener process has a fixed variance; by multiplying a distribution
with a fixed variance by a constant, one can allow the errors to have any variance.

The key difference in Equations 1 and 2 is the estimation of the autoregressive
relationships \( A_\Delta \) and \( A \), which are conceptually related but not equivalent. The parameter
\( A_\Delta \) tells a researcher about the autoregressive relationship for one sampling rate (e.g.,
“this is the autoregressive relationship for daily measurements”). The parameter \( A \), being
a continuous time value, describes the expected relationships for all possible lagged relationships (within the limits of the smallest and largest intervals covered by the data). These parameters are related through the equation

\[ A_\Delta = e^{A\Delta}, \]

(3)

where \( e \) is the symbol for exponent. In many applications, \( A_\Delta \) is expected to range between 0 and 1, the equivalent for \( A \) is to range from \(- \infty \) to 0; so a discrete time autoregressive relationship near 1 will be equal to a continuous time autoregressive relationship that is a negative number approaching zero².

As an example, let’s say that \( A_\Delta \) is equal to 0.9 and that measurements on our construct have been made every half-hour, then rewriting Equation 3 to solve for \( A \)

\[ A = \frac{\ln(A_\Delta)}{\Delta} = \frac{\ln(0.9)}{0.5} = -0.211. \]

(4)

Note that \( \ln \) is the symbol for natural log (log base \( e \)). One way to interpret \( A \) is to solve for the discrete time effect (\( A_\Delta \)) for many possible sampling intervals (\( \Delta_\Delta \)), as in Figure 19.7a. In this figure, the line consists of the discrete time autoregressive values (y-axis) for a range of sampling intervals (x-axis). That is, this continuous time parameter can be interpreted as giving information as to what the discrete time parameter would have been had we measured every 15 minutes, every 101 minutes, or any other possible sampling interval. The circle, on the other hand, is the information conveyed by \( A_\Delta \), the discrete time parameter. It is not that the discrete and continuous time approaches are different models, but rather, they convey the information in one’s data in very different ways, with the information offered by the discrete time approach being more limited. Unfortunately, it has been shown that fitting the discrete model and converting to the continuous time model parameter(s) (called the “indirect method”), as was done in this example, can lead to serious problems in parameter estimation that can result in misleading inferences (Hamerle, Nagl, & Singer, 1991). This means that to get the continuous time parameter, one needs to fit the continuous time model directly to data.

The application of Equation 2 to data can be accomplished in structural equation modeling software (Oud, 2007; Oud & Jansen, 2000). The exact fitting of this equation, unfortunately, requires a program that is adept with nonlinear constraints and can take the exponential of a matrix (e.g., Mx or OpenMx; Neale, Boker, Xie, & Maes, 2003; Boker et al., 2011). The Approximate Discrete Model is an approximation of Equation 2 that has been shown to provide reasonable approximations for the continuous time parameter \( A \)—much better than fitting the discrete time model and calculating \( A \) from \( A_\Delta \) (Bergstrom, 1988; Oud, 2010)³. The approximate discrete model fits a model that is partially described in the SEM in Figure 19.5e. On first inspection this figure would seem to be impossible to fit with SEM software. The single-headed arrow of an observed variable to itself is not an error variance but, rather, an instantaneous path from the observed variable to itself (recursive path). This oddity is possible through a set of linear
Dynamical Systems and Models of Continuous Time

constraints, constraints that define what the paths between observations and recursive paths should be equal to, given a value of $A$. In setting the linear constraints properly, one can achieve an estimate of $A$ (continuous time) rather than $A_A$ (discrete time).

Appendix A provides code for the software program R (using the package OpenMx; R, 2012; Boker et al., 2011). Examining this code, the reader will see that it sets up an SEM, as in Figure 19.5e. More importantly are the list of constraints; these constraints relate the paths in the figure back to the parameter $A$. Constraints on the instantaneous paths (recursive paths) and lagged paths (paths between different time) are:

$$A_{\text{instantaneous}} = \frac{1}{2} A_{\text{approx}} \Delta$$

$$A_{\text{lagged}} = 1 + \frac{1}{2} A_{\text{approx}} \Delta,$$

where $A_{\text{approx}}$ is a reasonable approximation to the continuous time parameter $A$. As all of the paths have constrained relationships with the parameter $A_{\text{approx}}$, in a case such as the one being illustrated the continuous time structural equation model (Fig. 19.5e) will be based on the same number of parameters as the discrete time model (Fig. 19.5a)—that is, the degrees of freedom will be equivalent for the two models. Additional discussion of these constraints, as well as the constraints placed on the parameter $G$, is available in Oud (2007, 2010).

It should be noted that the model presented in this section is the simplest of the possible models. This model can be made substantially more realistic through the inclusion of a measurement equation (relating latent variables rather than observed variables), inclusion of mean structure related to time (e.g., a developmental trajectory), inclusion of changes in relation to other time-varying variables, autocorrelation matrices $A$ that vary with time (a changing dependence on previous observations), and random effects for different individuals. The equations and presentation also focused on the examination of only a single variable; however, Equation 2 is easily altered to accommodate multiple variables and the relationships among those variables (discussed in discrete time as cross-lags), by changing the parameters and variables to matrices and changing the “1” in Equation 6 to an identity matrix. Example of these additions are discussed in several articles (e.g., Delsing, Oud, & De Bruyn, 2005; Delsing & Oud, 2008; Toharudin, Oud, & Billiet, 2008).
What is perhaps most interesting about the continuous time models is the additional understanding it can convey regarding one’s data. Figure 19.7b and 19.7c show some examples of discrete time parameter estimates ($A_\Delta, y$-axis) for a particular sampling interval ($x$-axis) based on three coupled constructs. The lines in Figure 19.7b represent the effects of constructs on themselves (two variables shown of three); the lines Figure 19.7c represent the effects of constructs on each other (two relationships shown of six). Using a discrete time analysis, one would only examine a singular value on the $x$-axis. Consequently two researchers collecting daily and weekly measurements (vertical lines) would find the effects of constructs on themselves (Figure 19.7b) to be positive and negative, respectively. In addition, the first researcher would say that there is a positive relationship between constructs, whereas the second would insist there is a negative relationship (solid line, 19.7c). Both would report their results, and the literature would be divided: Is the effect of A on B positive or negative? To add to the confusion, there would be similarities in their results (dashed line, 19.7c). These misunderstandings could be reconciled using a continuous time model, which could be used to produce figures such as Figure 19.7b and 19.7c and, therefore, give an impression as to the effect of one variable on the other for many possible sampling intervals.

Second-Order Differential Equation Model

Another continuous time model that one could consider would be a second-order differential equation model; this chapter specifically examines the second-order model that corresponds to a damped linear oscillator—that of a pendulum. This model is
Many simple pendulums vary around a point called the *equilibrium*; this is the point where a pendulum would come to rest given friction. The equation for this second-order differential equation is

\[
\frac{d^2x}{dt^2} = \eta x + \zeta \frac{dx}{dt},
\]

(7)

where \(\frac{d^2x}{dt^2}\), \(\frac{dx}{dt}\), and \(x\) are the second, first, and zeroth derivatives (acceleration, speed, and observed score) of the construct, \(\eta\) is related to the frequency of oscillation, and \(\zeta\) is related to the amount damping. In this equation, it is assumed that the equilibrium is constant, and has been set to zero. The parameter \(\eta\) is negative for a system that oscillates, such that when the construct score is high, there is a large negative acceleration; that is, if a person’s construct were to get far from their equilibrium there is an acceleration that will change their speed so that they start moving back toward equilibrium. If \(\eta\) is small then this restorative acceleration will be small and it will take a long time for the person to return to their equilibrium (low frequency), whereas a large negative number would provide a large restorative acceleration (high frequency). Over time—the addition of external forces on the pendulum can lead the pendulum to increase or decrease how far it swings (its amplitude). Changes in amplitude, regardless of whether they increase or decrease the amplitude, are called *damping*. Increases or decreases in the amplitude of the pendulum are conveyed in the \(\zeta\) parameter, with positive values corresponding to an increase in amplitude and negative values corresponding to a decrease in amplitude.

Most constructs are unlikely to change with the perfect oscillations expected of a pendulum. This, however, is not a requirement of the damped linear oscillator model if it is fit as a differential equation rather than using nonlinear estimation of a function such as sine. The differential equation only expresses a relationship between derivatives, stating that the distance a construct is from the equilibrium is related to the amount and direction of its acceleration. So, although this model matches the movements of a pendulum, it does not require the trajectories produced over time to conform to perfect oscillations. It is the ideas of equilibrium and restorative forces that make this model an interesting way that self-regulation could be conceptualized and modeled. The second-order differential equation with negative damping also conforms to the idea of a point attractor. Figure 19.8 shows two time series: one that corresponds to a pendulum-like oscillation and one that does not have perfect oscillation (left and right columns, respectively). The rows show the construct with respect to time (top) and plots of the relationships between derivatives (middle and bottom). Even when there are large...
departures from a pendulum-like oscillation, the relationships among derivatives remain similar to those of a pendulum.

Rather than solve the second-order differential equation in the manner done with the first-order differential equation, this section considers another option using SEM. In this approach, latent estimates of derivatives are estimated from observed data as described by Boker et al. (Latent Differential Equation Modeling; 2004). The second-order differential equation model is then fit by examining the paths between latent derivatives. The specification of this model does not require series of constraints, as is the case with the approximate discrete model. However, this model does require the time series data to be formatted in a specific way.

The data format that is necessary is called and embedded matrix—a concept from the state space literature. For our current purposes, we are interested in reconstructing a specific system (the damped linear oscillator model), so this treatment of embedded matrices can be relatively short. The key element to an embedded matrix is the number of embedding dimensions—the number of dimensions used to reconstruct a system (Takens, 1981). One way to think about the embedding dimension is that it will determine the number of observations used to estimate the moment-to-moment derivatives of a time series. As we are interested in a second-order model (i.e., a model with acceleration), we need to be able to estimate not just a straight line from data but also the curvature; we know that to estimate a curved line a minimum of three observations is required (four, if one wants to allow for error).

For the second-order model, this requirement of three or four observations to estimate the second derivative sets the lower bound for the embedding dimension—the minimum embedding that will faithfully reconstruct the dynamics of this system. However, how high the number of dimensions should be above the minimum is selected by the researcher and is motivated by two diametrically opposing goals. If one thinks about the

Figure 19.8 Plots of a pendulum-like oscillation (left column) and an oscillator with random disturbances (right column). The top row shows the plot of the construct over time, the subsequent rows show the relationships between estimates of the zeroth and first derivative with estimates of the second derivative.
embedding dimension as the number of observations that will be used to estimate any one derivative, then one can imagine using a lesser or greater number of observations. If using a lesser number of observations, then estimates will be more influenced by errors in the data and will be further (on average) from their true values than if one were to estimate the derivatives using a larger number of observations (i.e., higher variance in estimation). On the other hand, using a large number of observations, one will begin to average over true change of interest—that is, the true change variance will be reduced. The selection of the embedding dimension in other literatures is primarily motivated by the reconstruction of the system. In psychology and other social sciences, the selection of embedding dimensions must be considered in terms of how quickly the true system of interest is changing and selecting a dimension that strikes a balance between the amount of error reduction that occurs (by using a large embedding dimension) and maximizing the amount of true variance examined (by using a smaller embedding dimension).

Once the embedding dimension is selected, the physical creation of an embedded matrix is straightforward. To create an embedded matrix from a time series \( x \), where \( x \) has values \( x_1, x_2, \ldots, x_t \), one must rearrange multiple copies of the series into a matrix where adjacent columns are offset in time. For example, for an embedded matrix with embedding dimension four, we would produce a four column matrix

\[
X = \begin{bmatrix}
  x_1 & x_2 & x_3 & x_4 \\
  x_2 & x_3 & x_4 & x_5 \\
  \vdots & \vdots & \vdots & \vdots \\
  x_{t-3} & x_{t-2} & x_{t-1} & x_t
\end{bmatrix}
\]

(8)

This matrix will be entered into SEM software as if it consists of four observed variables, or however many columns one has selected as the embedding dimension. Interested readers can read more introduction to the selection of embedding dimensions and creation of an embedded matrix in Boker et al. (2004) and Deboeck (2011).

Figure 19.9 shows the SEM that will fit the model in Equation 7 to the embedded matrix. Each of the columns of the embedded matrix correspond to one of the observed variables. The paths from the latent variables to the observed variables, like in latent growth curve modeling, are all fixed so that the meaning of the latent variables is defined. The key difference here, in comparison to latent growth curve modeling, is the embedded matrix that treats the data as if we wish to specify lots of little growth curves along the entire length of the time series. The specification of the paths is not quite the same as in latent growth curve modeling—particularly for the second derivative and higher order derivatives—but still bears a resemblance in its estimation of the score of a construct at some time (intercept/zeroth derivative), estimation of how that construct is changing (slope/first derivative), and estimation of how the speed of the scores is accelerating or decelerating (curvature/second derivative). The syntax provided in Appendix B gives the
path values for an embedding dimension four and can be altered for any embedding dimension (see, Boker et al., 2004).6

As with the first-order differential equation model, this section has primarily focused on the simplest of cases: fitting a second-order differential equation model to a single time series. Not discussed here are topics such as how to set the equilibrium to zero, how to include multiple measures of the same construct, or how to analyze the data from multiple individuals. Many of these topics are discussed in the Boker et al. (2004) article and another chapter by Deboeck (2011), as well as examples that have been published using Latent Differential Equation Modeling (Bisconti, Bergeman, & Boker, 2004, 2006; Boker & Laurenceau, 2006; Boker, Leibenluft, Deboeck, Virk, & Postolache, 2008). There are also methods available for producing observed derivative estimates using equally spaced observations (Local Linear Approximation; Boker & Nesselroade, 2002; Boker & Graham, 1998), and unequally spaced observations (Generalized Orthogonal Local Derivative Estimates; Deboeck, 2010). Like the first-order differential equation model, we can also think about the coupling of multiple oscillators. Boker and Lauranceau (2006) have worked on several examples of coupled oscillators, examining the intimacy and disclosure patterns of husbands and wives; two coupled pendulums can produce remarkably complex change over time (e.g., Figure 19.1). The work by Boker and Lauranceau also has demonstrated some of the more nuanced questions that can be asked about coupling through this model, such as: Is it the husband’s level of intimacy that affects his wife, or is it the change in level of his intimacy that affects his wife?

Figure 19.9 Structural equation model using latent differential equation modeling to fit a damped linear oscillator model to an embedded matrix of observed values.

Conclusions & Future Directions

This chapter has given a brief introduction to dynamical systems, continuous time models, and methods for applying these ideas to data. The methods for the first-and second-order differential equation models, in particular, demonstrate some relatively new ways to address questions about change in data sets consisting of time series with significant proportions of measurement error. These models are the work of fusing
dynamical systems with statistics, taking into account the data constraints often 
experienced in the social sciences. These tools continue to expand, and the next decade is 
likely to continue to see the emergence of better methods as well as new ideas for other 
differential equation models that may be widely applicable to the social sciences.

This chapter did not go into great detail on many topics but has aimed to introduce key 
terms and ideas to give direction for further reading. Although two methods were 
mentioned in this article, there are a variety of other methods being used to fit dynamical 
systems to data. For example there are methods that directly apply equations to observed 
data such as in Dynamical Causal Modeling (Friston, Harrison, & Penny, 2003), direct 
comparison of observed and expected matrices (Deboeck & Boker, 2010), and methods 
that through iterative prediction try to obtain better estimates of observed values and 
parameter estimates such as Kalman filtering (Chow, Ferrer, & Nesselroade, 2007). The 
two methods introduced in this chapter were selected because they have been 
implemented in SEM software, making them (at present) a bit more accessible than other 
methods. No doubt this will change as more researchers ask questions about the 
relationships between the current state of constructs and how they are changing.

Dynamical Systems Theory, however, does come at a cost of re-evaluating some of the 
ways that data are routinely collected and analyzed. Related to dynamical systems is an 
area called Ergodic Theory. The mathematics in this area have highlighted that unless 
that all people have the same dynamical relationships—that is, all people change 
according to the same rules—inferences made using interindividual analyses are unlikely 
to be informative about any particular individual (Molenaar, 2004). This suggests that if 
one wishes to discuss the individual, which seems pertinent to much of psychology, 
models will have to be applied within individual before looking at interindividual 
differences. Intraindividual models will require many intraindividual measurements.

Despite the problems of data collection and the difficulty of fitting novel models, 
dynamical systems have become a point attractor. Researchers caught in this attractor 
are addressing new questions about how people change, sometimes analyzing old data 
with new methods and coming to a new understanding of those data. And although in the 
past dynamical systems seemed to require an impossibly large number of observations 
(many articles would discuss the need for thousands of observations), the combination of 
ew new methods for ambulatory assessments and statistical methods developed for shorter 
time series will no doubt continue to improve our ability to glean information from the 
seemingly random, complex variation of individuals.

Appendix A: Approximate Discrete Model

The following syntax applies the approximate discrete model to a matrix named “data” 
with $N$ rows (one row per subject) and 5 columns. The syntax is written for the statistical
program R (2012). Users will need to install the R package OpenMx (Boker et al., 2011) prior to running this syntax. This syntax is also available on the website of the author.

Mant comments have been placed in the code, following the # character. The observations in this example were spaced to occur at times 0, 1, 3, 6 and 10; that is, there are different lags between each pair of observations. These lags can be changed by altering the “Delta” matrices in the code below. The model summary provides an estimate of “Aapprox” which is the approximation of the continuous time parameter $A$. This code uses a raw data matrix and Full Information Maximum Likelihood estimation.
Dynamical Systems and Models of Continuous Time

```r
rm(list=ls()) # clear workspace
library(openmx) # load OpenMx package
colnames(data) <- paste('x',c(1:5),sep='') # assign names to data matrix columns
manifestvariables <- colnames(data)

# begin model, provide raw data matrix
ADMSmodel <- mxModel("ADMS",mxData(data,type="raw"),
    # set up matrices with lag information
    mxMatrix(type="Full",nrow=1,ncol=1,free=FALSE,values=1,name="Delta"),
    mxMatrix(type="Full",nrow=1,ncol=1,free=FALSE,values=2,name="Delta2"),
    mxMatrix(type="Full",nrow=1,ncol=1,free=FALSE,values=3,name="Delta3"),
    mxMatrix(type="Full",nrow=1,ncol=1,free=FALSE,values=4,name="Delta4"),
    # create Asymmetric matrix
    mxMatrix(type="Full",nrow=5,ncol=5,byrow=TRUE,name="Asymmetric"),
    # tell OpenMx with values will be estimated
    free=c(FALSE,FALSE,FALSE,FALSE,FALSE,
          TRUE,TRUE,FALSE,FALSE,FALSE,
          FALSE,TRUE,TRUE,FALSE,FALSE,
          FALSE,FALSE,TRUE,FALSE,FALSE),
    # give estimator values unique names
    labels=c("NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA","NA"),
    mxMatrix(type="Full",nrow=1,ncol=1,free=FALSE,values=1,labels=c("Approx"),name="anmatrix"),
    mxConstraint(.5*Delta1*Approx=Asym12),
    mxConstraint(.5*Delta2*Approx=Asym13),
    mxConstraint(.5*Delta3*Approx=Asym14),
    mxConstraint(.5*Delta4*Approx=Asym15),
    mxConstraint(.5*Delta1*Approx=Asym21),
    mxConstraint(.5*Delta2*Approx=Asym22),
    mxConstraint(.5*Delta3*Approx=Asym23),
    mxConstraint(.5*Delta4*Approx=Asym24),
    # create Symmetric matrix
    mxMatrix(type="Full",nrow=5,ncol=5,byrow=TRUE,name="Symmetric"),
    free=c(TRUE,FALSE,FALSE,FALSE,FALSE,
          FALSE,TRUE,FALSE,FALSE,FALSE,
          FALSE,FALSE,TRUE,FALSE,FALSE,
          FALSE,FALSE,FALSE,TRUE,FALSE),
    labels=c("var1","NA","NA","NA","NA",
             "NA","var2","NA","NA","NA",
             "NA","NA","var3","NA","NA",
             "NA","NA","NA","var4","NA",
             "NA","NA","NA","NA","var5"),
    mxConstraint(Delta1*G*G=var2),
    mxConstraint(Delta2*G*G=var3),
    mxConstraint(Delta3*G*G=var4),
    mxConstraint(Delta4*G*G=var5),
    # create other matrices needed to calculate covariance
    mxMatrix(type="Full",nrow=1,ncol=5,free=TRUE,labels=paste("M",c(1:5),sep=""),
             name="Means"),
    # calculate expected covariance
    mxMatrix(type="Symmetric",nrow=5,ncol=5,free=TRUE,
             labels=paste("C",c(1:5),sep=""),
             name="Covar"),
    # create matrix of means
    mxMatrix(type="Full",nrow=5,ncol=1,free=TRUE,values=apply(data,2,mean),
             name="Means"),
    # identify optimization objective
    mxAlgebra(1,Covar*Means+Covar,labels="Objective"),
    # finished writing model
    model <- mxRun(ADMSmodel) # run model, save output as 'model'
    summary(model) # get summary of output
```

Page 24 of 31

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2018. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: Montclair State University; date: 18 December 2018
Appendix B: Latent Differential Equation Modeling

The following syntax used latent differential equation modeling to apply the damped linear oscillator model to a single time series named “time-series.” The syntax is written for the statistical program R (2012). Users will need to install the R package OpenMx (Boker et al., 2011) prior to running this syntax. This syntax is also available on the website of the author. Many comments have been placed in the code, following the # character. The model summary provides estimates of the frequency (η) and damping (ζ) parameters. This code uses Full Information Maximum Likelihood estimation. A function “Embed” is provided to embed the time series.
Dynamical Systems and Models of Continuous Time

```r
rm(list=ls()) # clear workspace
library(OpenMx) # load OpenMx package

# create function to embed data
# X=time series, Embedding dimension
Embed <- function(x,E) {
  len <- length(x)
  out <- x[1:(len-E+1)]
  for(i in 1:E) { out <- cbind(out,x[(i+1-1):(len-E+1)]) }
  return(out)
}

data <- Embed(timeseries,4) # embed timeseries
colnames(data) <- paste("xy",c(1:4),sep="") # add names to embedded matrix

lag <- 1 # time between equally spaced observations
# manifest and latent variable names
ObsVar <- paste("y",c(1:4),sep="")
MatNames <- c(ObsVar,c("zeroth","first","second"))

# create asymmetric matrix
A <- mxMatrix(type='Full',nrow=length(MatNames),ncol=length(MatNames),
              free=FALSE, name="A")
Avalues[1:4,5] <- c(2,1,1,1)
Avalues[1:4,4] <- c(-1.5,-0.5,0.5,1.5)*lag
Avalues[1:4,7] <- c(0.125,0.125,0.125,0.125)*(lag^2)
Alabels[7,5] <- "Eta"
Afree[7,5] <- TRUE
Alabels[7,6] <- "Setz"
Afree[7,6] <- TRUE

# create symmetric matrix
S <- mxMatrix(type='Symm',nrow=length(MatNames),ncol=length(MatNames),
              free=FALSE, name="S")
Slabels[7,6] <- c(paste("EY",c(1:length(ObsVar)),sep=""),
                 "zeroth","first","second")
Slabels[7,6] <- "CovFirstZeroth";
Slabels[6,5] <- "CovFirstZeroth"

# other matrices needed for covariance algebra, mean estimation
I <- mxMatrix(type='Iden',nrow=length(MatNames),name="I")
F <- mxMatrix(type='Full',nrow=length(ObsVar),ncol=length(MatNames),
              free=FALSE, name="F")
Fvalues[1:4,1:4] <- 1
M <- mxMatrix("Full",nrow=1,ncol=length(MatNames),name="M",
              labels=c("MU","MU","MU","MU",rep(NA,3)),
              free=rep(TRUE,4),rep(FALSE,3))

# create LLO model
DLOmodel <- mxModel("Model",A, S, I, F, M, # include matrices
# covariance algebra
mxAlgebra(F%*%solve(I-A)%*%F%*%solve(I-A))%*%F, name="ECov",dimnames=list(ObsVar,ObsVar)), # mean algebra
mxAlgebra(t(F%*%solve(I-A)%*%M), name="ExpM",dimnames=list(NA,ObsVar)),
# provide data
mxData(data,type="raw"), # provide optimization objective
mxFIMLObjective("ECov","ExpM")
) # finished writing model
DLOout <- mxRun(DLOmodel) # run model, save output as 'DLOout'
supply(DLOout) # get summary of output
```

PRINTED FROM OXFORD HANDBOOKS ONLINE (www.oxfordhandbooks.com). © Oxford University Press, 2018. All Rights Reserved. Under the terms of the licence agreement, an individual user may print out a PDF of a single chapter of a title in Oxford Handbooks Online for personal use (for details see Privacy Policy and Legal Notice).

Subscriber: Montclair State University; date: 18 December 2018
Author Note

Correspondence concerning this chapter can be addressed to Pascal R. Deboeck, Ph.D., University of Kansas, Department of Psychology, 1415 Jayhawk Blvd., Lawrence, Kansas 66045. http://people.ku.edu/~pascal/.

References


Dynamical Systems and Models of Continuous Time


Notes:

1. Clearly, this only works if all of the differing intervals have a common multiple.

2. The relationship between Equation 1 and Equation 2—that is, how to solve for Equation 3—is addressed in many introductory resources on stochastic differential equations(e.g., Björk, 2009; Phillies, 2000; van Kampen, 2007). In the statistical physics literature the Wiener process is also called Brownian motion. Langevin’s Equation, for example, describes the motion of a Brownian particle and is very similar to the equations presented with velocity $v$ replaced by position $x$ (Langevin, 1908). This same first-order system has also been solved assuming that $A$ is an $n \times n$ matrix rather than a constant, where $n$ represents the number of variables being analyzed (Bergstrom, 1990).

3. Essentially, the approximate discrete model is the result of using a trapezoidal rule, rather than truly integrating the model.

4. There is code published for the approximate discrete model for LISREL in Oud (2007).

5. With positive damping (increasing amplitude), it conforms to the idea of a point repeller, and with no damping it conforms to the idea of a cyclic attractor.

6. Appendix B also provides a function to embed a time series. Mplus code is provided on the website of the author to fit the SEM.
Pascal R. Deboeck
Pascal R. Deboeck, Department of Psychology, University of Kansas, Lawrence, KS